Chapter 4
The Three-Dimensional Structure of Proteins

SUMMARY

Section 4.1

- Proteins are made up of long chains of amino acids. The composition and order of the amino acids are critical to the protein function.
- For any native protein, there is one, or at most a few, three dimensional structures that functions correctly
- Protein structure can be classified into primary, secondary, tertiary, and quaternary structure
- Primary structure is the order of the amino acids. Secondary structure is characterized by a repetitive organization of the peptide backbone. Tertiary structure refers to the complete three-dimensional structure of the protein. Quaternary structure describes a protein that has multiple polypeptide chains.

Section 4.2

- The primary structure of a protein determines the other levels of structure
- A single amino acid substitution can give rise to a malfunctioning protein, as in the case with sickle-cell anemia.

Section 4.3

- Secondary structures are based upon periodic structures of the peptide backbone.
- The most common secondary structures are the α-Helix and β-Sheet.
- Native proteins may have combinations of various secondary structures
- Regions of secondary structures can be combined to form supersecondary structures, motifs, and domains
- One of the most common structures is the triple-helix of collagen, the protein that makes up the bulk of connective tissue.

Section 4.4

- Tertiary structure is the complete three-dimensional arrangement of all the atoms in a protein.
- The tertiary structure of proteins is maintained by different types of covalent and non-covalent interactions.
- Hydrogen bonding occurs between atoms on the peptide backbone as well as atoms in the side chains.
- Electrostatic attractions between positively charged side chains and negatively side chains are also important
- The tertiary structure of proteins is determined by the techniques of X-ray diffraction and nuclear magnetic resonance.
- Myoglobin, the first protein to have its tertiary structure determined, is a globular protein for oxygen storage.
- Myoglobin is a single polypeptide chain with 153 amino acids, 8 α-helical regions, and a prosthetic group called heme.
- The heme has a coordinated iron ion at the center that binds to oxygen.

- Proteins can be denatured by heat, pH, and chemicals. Denaturation causes the protein to lose its native tertiary structure.
- Some types of denaturation can be reversed, while others are permanent.

Section 4.5

- Quaternary structure is the final level of protein structure and pertains to those proteins that consist of multiple polypeptide chains. Each chain is called a subunit.
- subunits interact with each other through non-covalent interactions.
- some proteins with multiple subunits are allosteric, which means that the subunits interact such that binding of a ligand to one subunit affects the binding of ligands to other subunits.
- Hemoglobin is a classic example of protein quaternary structure. The protein has 4 subunits, two α-chains and two β-chains, and it exhibits positive cooperativity. Binding of oxygen to one subunit makes it easier for oxygen to bind to other subunits.
- Hemoglobin's affinity for oxygen is controlled by several factors including oxygen pressure and pH. When the pH drops or when oxygen pressure is low, hemoglobin tends to release more oxygen to the tissues. When the pH is high and oxygen is plentiful, such as at the lung-blood interface, hemoglobin binds oxygen.
- Hemoglobin is bound to 2,3-bisphosphoglycerate, which acts as a bridge between the 4 subunits. In the absence of 2,3- bisphosphoglycerate, hemoglobin is not allosteric and behaves like myoglobin.

Section 4.6

- Using the power of computers, we can now predict the tertiary structure of a protein if we know its amino acid sequence.
- A great deal of information regarding protein structure and sequences can be found on the World Wide Web.
- Chaperones are proteins that help another protein attain the correct native conformation.
- There is a specific chaperone for the formation of hemoglobin.
- Protein folding is critical to the proper function of a protein. There are diseases caused by misfolded proteins. One of the most infamous is a disease caused by a misfolded protein called a prion. Misfolded prions cause spongiform encephalopathy, otherwise known as Mad Cow Disease in the dairy industry or Creutzfeld-Jakob disease in humans. Other important diseases that involve misfolded proteins are Alzheimers, Parkinsons, and Huntingtons.

LECTURE NOTES

This chapter presents details on the general structure of proteins. The hierarchical nature of the four levels of structure in proteins leading to the final three-dimensional shape of the protein helps students grasp these all-important concepts.

This chapter requires two, and more likely, three lectures. The first will likely be devoted to primary and secondary structures and the forces that govern their formation. The second lecture might cover tertiary and quaternary structures, along with representative examples. A third lecture might involve more detailed discussion of higher-order structure, using the properties of hemoglobin and myoglobin as contrasting examples, linking structure with function.

LECTURE OUTLINE

I. Protein structure linked to function
    A. Native versus denatured conformations
    B. Levels of protein structure and the forces holding them together
        1. Primary – peptide bonds
        2. Secondary – hydrogen bonds
        3. Tertiary – interplay of side chains
        4. Quaternary – noncovalent interactions
II. Primary structure
III. Secondary structure
    A. Peptide planes, Φ and ψ angles
    B. The α-helix
        1. Position of hydrogen bonds
        2. Dimensions
        3. Side chain considerations
    C. The β-sheet
    D. Other structures
        1. Bulges and turns
        2. Supersecondary domains
        3. Collagen
        4. Fibrous vs. globular proteins
IV. Tertiary structure
    A. Use of X-ray crystallography and NMR studies
    B. Myoglobin as an example of protein structure
        1. Folding around heme
        2. Coordination of iron
    C. Denaturation and refolding
V. Folding predictions from sequence
    A. Bioinformatics and sequence homology
    B. Chaperones

VI. Quaternary structure
- A. Hemoglobin as an example of protein structure
- B. Conformational changes and hemoglobin function
  1. Positive cooperativity of O2 binding
  2. Bohr effect
  3. Effects of BPG binding

VII. Thermodynamics and protein folding
- A. Weak forces
- B. Importance of hydrophobic interactions

ANSWERS TO PROBLEMS

4.1 Protein Structure and Function

1. (a) (iii); (b) (i); (c) (iv); (d) (ii).
2. When a protein is denatured, the interactions that determine secondary, tertiary, and any quaternary structures are overcome by the presence of the denaturing agent. Only the primary structure remains intact.
3. The random portions of a protein do not contain structural motifs that are repeated within the protein, such as α-helix or β-pleated sheet, but three-dimensional features of these parts of the protein are repeated from one molecule to another. Thus, the term *random* is somewhat of a misnomer.

4.2 Primary Structure of Proteins

4. When a protein is covalently modified, its primary structure is changed. The primary structure determines the final three-dimensional structure of the protein. The modification disrupts the folding process.
5.

(a) Serine has a small side chain that can fit in any relatively polar environment.

(b) Tryptophan has the largest side chain of any of the common amino acids, and it tends to require a nonpolar environment.

(c) Lysine and arginine are both basic amino acids; exchanging one for the other would not affect the side-chain pK_a in a significant way. Similar reasoning applies to the substitution of a nonpolar isoleucine for a nonpolar leucine.

6. Glycine is frequently a conserved residue because its side chain is so small that it can fit into spaces that will not accommodate larger ones.
7. When alanine is replaced by isoleucine, there is not enough room in the native conformation for the larger side chain of the isoleucine. Consequently, there is a great enough change in the conformation of the protein that it loses activity. When glycine is substituted in turn for isoleucine, the presence of the smaller side chain leads to a restoration of the active conformation.
8. Meat consists largely of animal proteins and fat. The temperatures involved in cooking meat are usually more than enough to denature the protein portion of the meat.

4.3 Secondary Structure of Proteins

9. Shape, solubility, and type of biological function (static, structural versus dynamic, catalytic).

10. The angles of the amide planes as they rotate about the α-carbon. The angles are both defined as zero when the two planes would be overlapping such that the carbonyl group of one contacts the N—H of the other.
11. A β-bulge is a common nonrepetitive irregularity found in antiparallel β-sheets. A misalignment occurs between strands of the β-sheet, causing one side to bow outward.
12. A reverse turn is a region of a polypeptide where the direction changes by about 180°. There are two kinds—those that contain proline and those that do not. See Figure 4.7 for examples.
13. The α-helix is not fully extended, and its hydrogen bonds are parallel to the protein fiber. The β-pleated sheet structure is almost fully extended, and its hydrogen bonds are perpendicular to the protein fiber.
14. The αα unit, the βαβ unit, the β-meander, the Greek key, and the β-barrel.
15. The geometry of the proline residue is such that it does not fit into the α-helix, but it does fit exactly for a reverse turn. See Figure 4.7c.
16. Glycine is the only residue small enough to fit at crucial points in the collagen triple helix.
17. The principal component of wool is the protein keratin, which is a classic example of α-helical structure. The principal component of silk is the protein fibroin, which is a classic example of β-pleated sheet structure. The statement is somewhat of an oversimplification, but it is fundamentally valid.
18. Wool, which consists largely of the protein keratin, shrinks because of its α-helical conformation. It can stretch and then shrink. Silk consists largely of the protein fibroin, which has the fully extended β-sheet conformation, with far less tendency to stretch or shrink.

4.4 Tertiary Structure of Proteins

19. See Figure 4.2 for a hydrogen bond that is part of the α-helix (secondary structure). See Figure 4.12 for a hydrogen bond that is part of tertiary structure (side-chain hydrogen bonding).
20. See Figure 4.12 for electrostatic interactions, such as might be seen between the side chains of lysine and aspartate.
21. See Figure 4.12 for an example of a disulfide bond.
22. See Figure 4.12 for an example of hydrophobic bonds.
23. *Configuration* refers to the position of groups due to covalent bonding. Examples include *cis* and *trans* isomers and optical isomers. *Conformation* refers to the positioning of groups in space due to rotation around single bonds. An example is the difference between the eclipsed and staggered conformations of ethane.
24. Five possible features limit possible protein configurations and conformations. (1) Although any one of 20 amino acids is possible at each position, only one is used, as dictated by the gene that codes for that protein. (2) Either a D- or an L-amino acid could be used at each position (except for glycine), but only L-amino acids are used. (3) The peptide group is planar, so that only *cis* and *trans* arrangements are observed. The *trans* form is more stable and is the one usually found in proteins. (4) The angles ϕ and ψ can theoretically take on any value from 0° to 360°, but some angles are not possible because of steric hindrance;

angles that are sterically allowed may not have stabilizing interactions, such as those in the α-helix. (5) The primary structure determines an optimum tertiary structure, according to the "second half of the genetic code."

25. Technically, collagen has quaternary structure because it has multiple polypeptide chains. However, most discussions of quaternary structure involve subunits of globular proteins, not fibrous ones like collagen. Many scientists consider the collagen triple helix to be an example of a secondary structure.

4.5 Quaternary Structure of Proteins

26. *Similarities:* both contain a heme group; both are oxygen binding; secondary structure is primarily α-helix. *Differences:* hemoglobin is a tetramer, while myoglobin is a monomer; oxygen binding to hemoglobin is cooperative, but noncooperative to myoglobin.
27. The crucial residues are histidines in both proteins.
28. Myoglobin's highest level of organization is tertiary. Hemoglobin's is quaternary.
29. The function of hemoglobin is oxygen transport; its sigmoidal binding curve reflects the fact that it can bind easily to oxygen at comparatively high pressures and release oxygen at lower pressures. The function of myoglobin is oxygen storage; as a result, it is easily saturated with oxygen at low pressures, as shown by its hyperbolic binding curve.
30. In the presence of H^+ and CO_2, both of which bind to hemoglobin, the oxygen-binding capacity of hemoglobin decreases.
31. In the absence of 2,3-*bis*phosphoglycerate, the binding of oxygen by hemoglobin resembles that of myoglobin, characterized by lack of cooperativity. 2,3-*Bis*phosphoglycerate binds at the center of the hemoglobin molecule, increases cooperativity, stabilizes the deoxy conformation of hemoglobin, and modulates the binding of oxygen so that it can easily be released in the capillaries.
32. Fetal hemoglobin binds oxygen more strongly than adult hemoglobin. See Figure 4.28.
33. Histidine 143 in a β-chain is replaced by a serine in a γ-chain.
34. Deoxygenated hemoglobin is a weaker acid (has a higher pK_a) than oxygenated hemoglobin. In other words, deoxygenated hemoglobin binds more strongly to H^+ than does oxygenated hemoglobin. The binding of H^+ (and of CO_2) to hemoglobin favors the change in quaternary structure to the deoxygenated form of hemoglobin.
35. The primary flaw in your friend's reasoning is a reversal of the definition of pH, which is pH = –log [H^+]. If the release or binding of hydrogen ion by hemoglobin were the primary factor in the Bohr effect, the pH changes would be the opposite of those actually observed. The response of hemoglobin to changes in pH is the central point. When the pH increases, the hydrogen ion concentration decreases, and vice versa.
36. The change of a histidine to a serine in the γ-chain removes a positively charged amino acid that could have interacted with BPG. Thus there are fewer salt bridges to break, so binding is easier than it is in a β-chain.
37. People with sickle-cell trait have some abnormal hemoglobin. At high altitudes, there is less oxygen, and the concentration of the deoxy form of the abnormal

hemoglobin increases. Less oxygen can be bound, causing the observed breathing difficulties.

38. In fetal hemoglobin, the subunit composition is $\alpha_2\gamma_2$ with replacement of the β-chains by the γ-chains. The sickle-cell mutation affects the β-chain, so the fetus homozygous for Hb S has normal fetal hemoglobin.
39. The relative oxygen affinities allow oxygen to be taken by the fetal cells from the maternal Hb.
40. Because people with sickle-cell disease are chronically anemic, some cells with fetal Hb are produced to help overcome the impaired oxygen delivery system.
41. The crystalline form changed because oxygen entered under the cover slip, transforming deoxyhemoglobin to oxyhemoglobin.

4.6 Protein Folding Dynamics

42. On the β chain of HbS there is a valine at position 6 where there is a glutamic acid on the normal form.
43. Valine is a hydrophobic amino acid. The valine is on the outside of the globular protein where the amino acids have to interact with water. The presence of the hydrophobic amino acid at that position causes hemoglobins to clump together via hydrophobic interactions. This clumping causes the cell deformation.
44. It appears that in the heterozygous genotype, there is an advantage in areas prone to malaria.
45. Hydroxyurea stimulates the bone marrow to produce fetal hemoglobin. Fetal hemoglobin has no β chain, so the effects of the interactions of those chains is reduced in the presence of increased HbF.
46. BCL11A is a protein that represses the production of HbF. It is relevant because if BCL11A can be repressed itself, then an adult would still produce HbF, which would help alleviate symptoms of sickle cell anemia.
47. The theory is the HbF evolved so that the fetal hemoglobin would have a higher affinity for oxygen than the mother's adult hemoglobin. This would favor transfer of oxygen from the mother to the fetus. Is there a downside to this in the adult? Nobody knows for sure as at present the focus is on how much better it would be to have HbF compared to HbS. However, we can think that having hemoglobin with an abnormally high affinity for oxygen may have a downside. For example, an athlete working his muscles very hard needs to have hemoglobin that will release the oxygen to the oxygen-starved tissues. Perhaps HbF would be too slow to accomplish this. In other words, one hypothesis would be that a person with all HbF would not make a very good aerobic athlete.
48. This level of sequence homology is marginal for use of comparative modeling. It is best to try that method, but then to compare the results with those obtained from the fold-recognition approach.
49. Protein folding is driven by many processes. The intuitive ones are the direct interactions of functional groups through covalent bonds, electrostatic attractions, and hydrogen bonds. These explain why parts of the protein are attracted to each other and why a protein would tend to adopt a shape making these interactions possible. However, much of the protein-folding process is also driven by an entropy effect. We refer to hydrophobic interactions as an explanation of

why nonpolar regions of the protein tend to cluster together, usually in the interior of the protein. However, in reality, it is not the interaction of nonpolar amino acids that drives this process. It is actually the increase in entropy of the solvent, water. When the hydrophobic regions of the protein are isolated to the interior, the water molecules surrounding the protein are more free to rotate and move in less restricted ways. Thus, what drives much of protein folding is not a ΔH change with the bonding of specific amino acids, but rather a ΔS increase of the solvent.

50. See the Protein Data Bank.
51. A chaperone is a protein that aids another protein in folding correctly and keeps it from associating with other proteins before it has reached its final, mature form.
52. A prion is a potentially infectious protein found in multiple forms in mammals, often concentrated in nervous tissue. It is an abnormal form of a normal cellular protein. It tends to form plaques that destroy the nervous tissue. Prions have been found to be transmissible across species.
53. A series of encephalopathies have been found to be caused by prions. In cows, the disease caused by prions is called bovine spongiform encephalopathy, or more commonly mad-cow disease. In sheep, the disease is called scrapie. In humans, it is called Creutzfeldt-Jakob disease.
54. The normal form of the prion protein has a higher α-helix content compared to the β-sheet content. The abnormal one has an increased β-sheet content.
55. Alzheimer's, Parkinson's, and Huntington's diseases are caused by accumulation of protein deposits from aggregates caused by misfolded proteins. This chapter also looked at prion diseases. When prions are misfolded they can cause spongiform encephalopathies, such as mad-cow disease, and the human form, Creutzfeldt-Jakob disease.
56. Protein aggregates form when there are exposed areas on a protein surface that are nonpolar. Proteins then stick together via these nonpolar regions causing the aggregates. An example is the prion disease in which an area of the normal molecule that should be an α-helix adopts a β-sheet conformation instead.
57. The root problem with the globin genes and potential issues with hemoglobin formation is based on the fact that there are two α-globin genes for every β-globin gene, yet to make hemoglobin they must combine in a 1:1 ratio. Thus, one theoretically possible solution would be if there were not a 2:1 ratio of these genes. Another issue is that the two genes are on different chromosomes. They are then most likely controlled separately. If the two genes were close together on one chromosome, then they could be controlled together by the same signal and produced in the correct amounts.
58. The sequence of the mutant prion that confers the most extreme sensitivity to conferring a prion disease is the substitution of the amino acid at position 129 to a methionine.
59. Prion diseases are transmissible, while other neurodegenerative diseases like Alzheimer's are not.
60. The spongiform encephalopathies that we know of have characteristics of both inherited diseases and transmissible diseases. On the one hand, animals can be infected by consuming meat or other tissues that are themselves carrying mutant prion proteins. As the example of the New Zealand sheep showed, even those

that are very susceptible to a prion disease can remain disease-free if they are never exposed. However, the predisposition to acquire a prion disease has a hereditary component as well. The prion protein has many known mutations, some of which render the individual very susceptible to the disease. These mutations can be tracked and they are passed along family lines.

61. The two enzymes associated with the disease are called β- secretase and γ-secretase.
62. Amyloid β and Tau are the two proteins that form destructive plaques. The former is formed from pieces cut from a precursor protein called Amyloid Precursor Protein.
63. Alzheimer's begins with the buildup of Aβ, which is cut from the APP. In the first step, the enzyme β-secretase cuts APP outside the cell membrane. Then the γ-secretase enzyme cuts the remaining portion of the APP inside the membrane, releasing Aβ.
64. β-secretase is naturally involved with the myelination of nerves.
65. Prion diseases have been linked to the immune system. It is believed that the prion proteins travel in the lymph system bound to lymphocytes and eventually arrive at the nervous tissue, where they begin to transform a normal cellular protein into an abnormal one (a prion).
66. Although there may be a strong genetic predisposition to acquire scrapie, that alone will not cause the disease. The disease must be started by ingesting a prion that already has the altered conformation, PrPsc.